

Research Article

Spatial Information Entropy and Its Application in the Degradation State Identification of Hydraulic Pump

Yu-kui Wang, Hong-ru Li, Bing Wang, and Bao-hua Xu

Shijiazhuang Mechanical Engineering College, No. 97, Heping West Road, Shijiazhuang 050003, China

Correspondence should be addressed to Yu-kui Wang; wktougao@163.com

Received 18 April 2015; Accepted 25 May 2015

Academic Editor: Michael Small

Copyright © 2015 Yu-kui Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The degradation state identification is a key step of the condition based maintenance of hydraulic pump. In this paper, spatial information entropy (SIE) as a novel degradation feature of pump is proposed based on the study of permutation entropy (PE) algorithm. The fundamental principle of SIE is introduced and contrasted with PE. Different parameters used in the calculation of SIE are discussed and meaningful conclusion is gained. The results of simulation analysis not only checked the rationality of SIE but also demonstrated the availability and superiority of adopting SIE as the degradation feature. Based on simulation analysis, SIE and PE are united and used as degradation feature vector of pump. FCM algorithm is employed to diagnose the degradation state of pump. The analysis results of practical signal testified the rationality and availability of the proposed method.

1. Introduction

As the heart of hydraulic system, the condition of hydraulic pump is important to the whole system. For the heavy load and complex working condition, pump is also the unit that usually fails in function [1]. As the failure will result in high industrial cost or even disaster, considerable attention has been paid to the condition monitoring and fault diagnosis of pump for a long time [2]. But most of the researches concentrate on the fault recognition and fault location [1, 3]. Researches on the degradation state identification of pump are rarely reported [4]. The method based on vibration signal analysis is extensively used for pump because of the advantages of fitness and validity [3].

With the development of maintenance theory and relative technologies, condition based maintenance is getting more and more attention [5–7]. A lot of work has been done about degradation state identification (DSI) as it is the fundament of condition based maintenance [8–11]. DSI has two important steps: one is extracting appropriate features which can reflect comprehensive degradation degree from the raw vibration signal; the other is building an effective intelligent model which is used to assess the state of equipment [12]. Proper feature extraction is the key step of DSI as it affects the

precision of final state identification. Feature extraction is the process of converting original data into relevant information of the equipment. Traditional features contain three categories: time domain, frequency domain, and time-frequency domain. They are sensitive for fault identification and are widely used in the fault diagnosis of mechanical equipment [13]. However, they also have defects; for example, their stability is not good enough; as a result, they cannot indicate the fault degree of equipment [14]. As the vibration signal displays strong nonlinear feature when the machinery is broken-down [15], a lot of nonlinear methods are applied to the vibration signal processing with the development of nonlinear theory, such as fractal dimension, approximate entropy, and sample entropy [16, 17]. Permutation entropy (PE) was proposed by Bandt et al. [18, 19] as a complexity indicator of time series. For its advantages of clear concept and simple calculation, it is widely used in the mutation detection of electroencephalogram, heart interbeat signal, and geomagnetic storms as well as mechanical signal [20–23]. From the researches mentioned above, it can be concluded that the research of degradation feature extraction acquired great progress in recent years.

In this paper, the degradation state identification of pump which reflects the relationship between its fault degree and

the degradation features is researched. And PE is used as the degradation feature of it. Based on the study of PE algorithm, we found that it only explored the sorting relationship of the elements in certain refactoring component on account of their value [18]. The spatial location information of each element in the original time series is not considered. In order to reveal the distribution information of refactoring component's elements in the original time series, the spatial information entropy (SIE) is put forward. And it is used as another degradation feature of pump. The relationship between degradation feature and fault degrees is nonlinear; therefore, the nonlinear model should be founded in order to obtain better result [4]. Furthermore, the fuzzy recognition method should be selected because the degradation process of pump is gradual and the description of its fault degree by degradation features is fuzzy and uncertain. The methods such as rough set theory, D-S evidence theory, and fuzzy set theory are widely used in the fuzzy system analysis and fuzzy information processing; their validity and superiority are demonstrated. However, they also have the shortage of requiring large amount prior knowledge and experiences [24]. Fuzzy C-means (FCM) clustering is a traditional but excellent algorithm for pattern recognition, with some advantages such as being simple, intuitive, and explicable [8]. In addition, FCM clustering transforms clustering into the nonlinear programming with constraints; therefore, just small amount training sample is needed in the clustering process and high precision ratio also can be obtained. It is widely used in fault diagnosis, image processing, and also other fields. In this paper, FCM is used to identify the degradation states of pump.

The remainder of this paper is organized as follows. In Section 2, the basic theory of PE is simply introduced. The proposed SIE scheme is detailed at first in Section 3, and the selection of its parameters is analyzed, based on which both PE and SIE are used to analyze the degradation simulation signal of pump, and the excellent performance of SIE is checked. In Section 4, the FCM clustering algorithm and the degradation state identification strategy are introduced. Section 5 introduces the pump degradation state experiment at first, and then the collected vibration signals are employed to evaluate the proposed method. Our conclusions are drawn in Section 6.

2. Permutation Entropy

Permutation entropy (PE) represents a new way to assess the complexity of nonlinear time series. PE has some advantages as compared to other entropy measures, since it is an ordinal measure. Indeed, PE decomposes the time series into a series of ordinal patterns describing the order relations between the present and a fixed number of equidistant past values at a given time. The mathematical theorem of permutation entropy was described in detail in [18, 19]. The PE of a time series $\{x(j), j = 1, 2, 3, \dots, N\}$ can be calculated as follows [18].

(1) Reconstruct the time series and its phase space can be obtained as follows:

$$\begin{bmatrix} x(1) & x(1+\tau) & \cdots & x(1+(m-1)\tau) \\ x(2) & x(2+\tau) & \cdots & x(2+(m-1)\tau) \\ x(i) & x(i+\tau) & \cdots & x(i+(m-1)\tau) \\ \cdots & \cdots & \cdots & \cdots \\ x(K) & x(K+\tau) & \cdots & x(K+(m-1)\tau) \end{bmatrix}, \quad (1)$$

where m is embedded dimension and τ is delay time, $i = 1, 2, \dots, K$, $K = N - (m-1)\tau$. Each row of the reconstructed matrix is a refactoring component.

(2) Extract the rank numbers of m elements and name them as their labels.

(3) Arrange m real values contained in refactoring component in increasing order.

(4) A symbol series can be obtained by using the labels instead of their real values.

(5) Extract the symbol series of all refactoring components by steps (1–4). Then, count the number of each existing symbol series and calculate their probability.

(6) If the probability are denoted by P_1, P_2, \dots, P_q , $\sum_{i=1}^q P_i = 1$, the PE of this time series can be defined as

$$H_p^O = - \sum_{i=1}^q P_i \ln P_i. \quad (2)$$

As each refactoring component contains m elements, the largest number of symbol series is $m!$. The maximum value of H_p^O can be obtained as $\ln(m!)$ when all the symbol series have the same probability distribution $P_i = 1/m!$. Therefore, the PE can be normalized with

$$0 \leq H_p = \frac{H_p^O}{\ln(m!)} \leq 1. \quad (3)$$

The value of H_p represents the randomness and complexity of the time series $\{x(j), j = 1, 2, 3, \dots, N\}$, and it also describes local order structure of the time series. The smallest possible value of H_p is zero, which means that the time series is very regular [18]. The largest possible value of H_p is 1, which can be realized when all symbol series have equal probability. When the pump has certain fault, the more serious the fault is, the more regular the components will exist in its vibration signal, which results in smaller randomness and lower complexity of the signal and also smaller PE of it. On the contrary, when the pump is normal, the randomness, complexity, and also PE of its signal reach their maximum. Therefore, PE can be used as degradation feature of pump to indicate its fault degree.

3. Spatial Information Entropy

PE can sensitively reflect the variation of the randomness and complexity of time series, so it is usually used to indicate the system dynamics mutations [18]. As it adopts the sorting variation of elements in certain refactoring component to

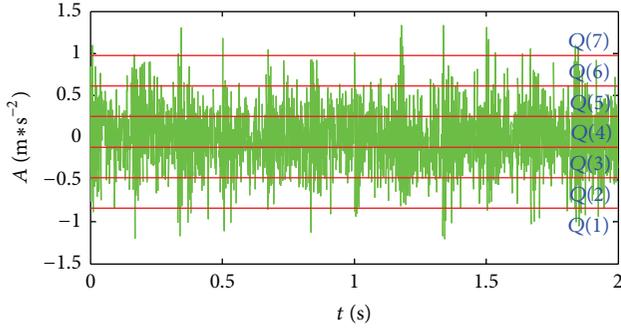


FIGURE 1: Partition example.

reflect the dynamics variation of the time series [19], the spatial location information of each element in the original time series is not considered. When the time series varies, the spatial location information of each element is sure to change; in order to reflect this change, the spatial information entropy is proposed; the feasibility and availability of using it to indicate the dynamics variation of time series are also analyzed in this section.

3.1. Principle. SIE of time series $\{x(j), j = 1, 2, 3, \dots, N\}$ can be calculated with the following steps.

(1) Find the maximum value $\max(x)$ and the minimum value $\min(x)$ of the time series, and divide the time series into s regions, which are denoted by $Q(1), Q(2), \dots, Q(s)$. Figure 1 shows the partition state of dividing a signal into 7 regions.

(2) Reconstruct phase space of the time series with parameters m as embedded dimension and τ as delay time; the same result in formula (1) can be obtained, and there are $N - (m - 1)\tau$ refactoring components in the reconstruction matrix.

(3) There are m elements in any refactoring component $[x(i), x(i + \tau), \dots, x(i + (m - 1)\tau)]$; each element belongs to a certain region in the original time series. For element x_i^d , $d = 1, 2, \dots, m$, its region label k_i^d can be extracted as follows.

If $x_i^d = \max(x)$, $k_i^d = s$. Otherwise, if formula (4) is validity, the conclusion of $x_i^d \in Q(k)$, $k = 1, 2, \dots, s$, and also $k_i^d = k$ can be gained:

$$\min(x) + D \times (k - 1) \leq x_i^d < \min(x) + D \times k, \quad (4)$$

$$D = \frac{[\max(x) - \min(x)]}{s}. \quad (5)$$

The spatial symbol series $[k_i^1, k_i^2, \dots, k_i^m]$ of refactoring component $[x(i), x(i + \tau), \dots, x(i + (m - 1)\tau)]$ can be gained by using the region labels to take place of real elements. The spatial symbol series of other refactoring components can be extracted in the same way. There are s regions that each element may belong to, so the maximum number of spatial symbol series is s^m .

(4) Assuming q ($q \leq s^m$) categories spatial symbol series are obtained, the number of every category spatial symbol series is counted; then, the probabilities of them are calculated and denoted by $P_K^1, P_K^2, \dots, P_K^q$, $\sum_{i=1}^q P_K^i = 1$, based on which

the SIE of the time series can be defined as the form of Shannon entropy with the following:

$$H_K^O = - \sum_{i=1}^q P_K^i \ln P_K^i. \quad (6)$$

The maximum value of H_K^O is $\ln(s^m)$ on condition that all the spatial symbol series have the same probability as $1/s^m$. Therefore, H_K^O can be normalized with

$$0 \leq H_K = \frac{H_K^O}{\ln(s^m)} \leq 1. \quad (7)$$

The value of H_K reflects the randomness and complexity of time series $\{x(j), j = 1, 2, 3, \dots, N\}$; the smaller value indicates the smaller randomness and the lower complexity of the time series. On the contrary, the larger H_K is, the more irregular the time series is. From the principle of SIE, we can conclude that it explores the distribution information of refactoring component's elements in the original time series and adopts their spatial distribution variation to reflect the dynamics variation of the time series. It is obvious that on basis of $s \geq m$ the possible categories of spatial symbol series s^m are significantly greater than $m!$ which are the possible categories of PE symbol series. Therefore, SIE should be able to reflect the dynamics variation of time series more excellently in detail than PE.

3.2. Parameters Selection. From the principle of SIE, it is clear that three parameters need to be considered and selected in the calculation of SIE, which are partition number s , embedding dimension m , and delay time τ . In this section, their influence is analyzed as to offer reference for the selection of them. In the research, the following simulation signal is designed to imitate the vibration signal of failure pump [25]:

$$x(t) = x_{is}(t) + x_{fs}(t) + n(t), \quad (8)$$

$$x_{is}(t) = e^{-60t_1} \sin(540\pi t), \quad (9)$$

$$x_{fs}(t) = e^{-30t_2} \sin(540\pi t),$$

$$t_1 = \text{mod}\left(t, \frac{1}{30}\right), \quad (10)$$

$$t_2 = \text{mod}\left(t, \frac{1}{6}\right).$$

There are two shocking components $x_{is}(t)$ and $x_{fs}(t)$ in the simulation signal, where $x_{is}(t)$ denotes the inherent shocking component of pump and $x_{fs}(t)$ denotes the shocking component generated by the fault of pump. The waveform of shocking components is shown in Figure 2. $n(t)$ denotes the noise in the signal, and its Signal to Noise Ratio (SNR) is -3 dB. The sampling frequency is 1024 Hz and sampling number is 2048. Figure 3 shows the time domain and frequency spectrum waveforms of $x(t)$.

In order to analyze the influence of the three parameters, set the ranges of them as $s \in [2, 3, \dots, 10]$, $m \in [1, 2, \dots, 7]$,

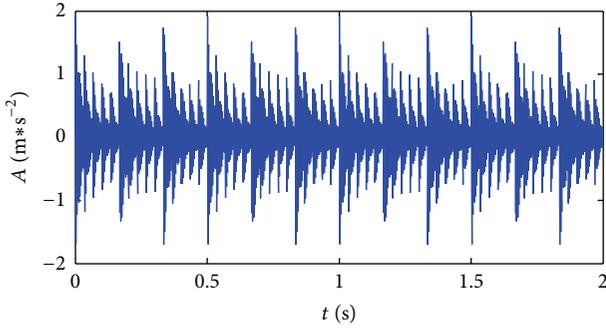


FIGURE 2: Waveform of shocking component.

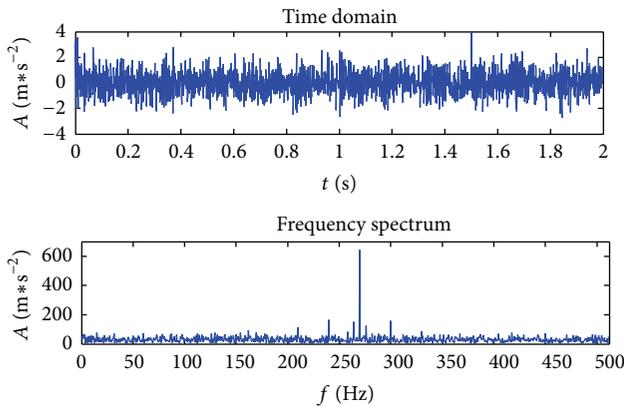
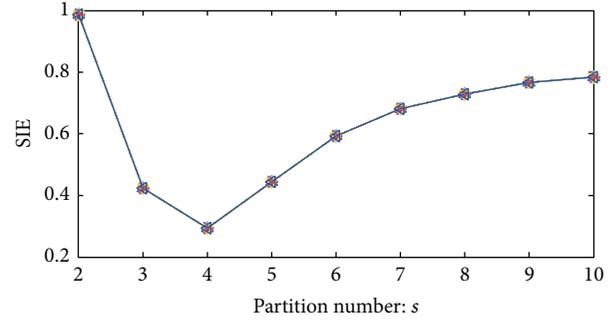


FIGURE 3: Waveform of simulation signal.

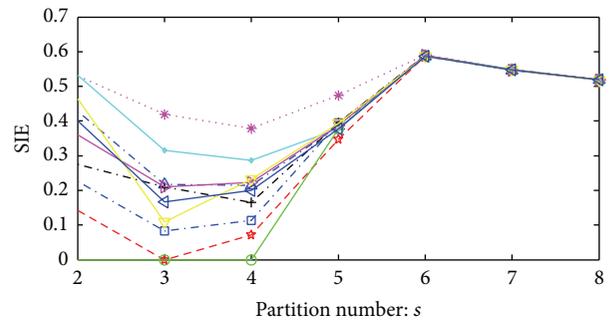
and $\tau \in [1, 2, \dots, 10]$ (in order to avoid data overflow, the range of s is 2–8 when $m = 7$). Calculate the SIE of $x(t)$ with all possible parameters combinations (s, m, τ). Figure 4 shows the results of parameters combination ($s, 1, \tau$) which means $m = 1$ and s and τ take all the values of their ranges. It can be seen that the values of SIE are different when s is changing; when s is determined, the value of SIE is invariant no matter what the value of τ is. When taking other parameters combinations, the SIEs are, respectively, calculated and analyzed; the following is the obtained conclusion: when $s \geq m - 1$, the variation of τ has little influence on SIE, if not, SIE changes greatly along with the variation of τ . Figure 5 is the waveforms of SIEs when $m = 7$. It is obvious that when $s < 6$, SIE changes greatly as the variation of τ ; otherwise, SIE is stable though τ is changing. The case is the same when m is 2–6, respectively; limited by space, the results of other parameters combinations are not shown here.

Ten SIEs can be obtained when s and m are determined as τ has ten different values. Variance of the ten SIEs is calculated to reflect the stability of the SIEs with certain (s, m). All the SIE variances with possible (s, m) are calculated and shown in Table 1. The results demonstrate that τ has little influence on SIE when $s \geq m - 1$. Therefore, we take $s \geq m$ as the principle for the selection of s and m , based on which τ can be randomly selected and we select $\tau = 3$.



Delay time: 1 (blue square), Delay time: 2 (blue triangle), Delay time: 3 (black plus), Delay time: 4 (magenta asterisk), Delay time: 5 (red star), Delay time: 6 (green circle), Delay time: 7 (cyan diamond), Delay time: 8 (magenta inverted triangle), Delay time: 9 (yellow inverted triangle), Delay time: 10 (blue left-pointing triangle)

FIGURE 4: SIEs when embedding dimension are 1.



Delay time: 1 (blue square), Delay time: 2 (blue triangle), Delay time: 3 (black plus), Delay time: 4 (magenta asterisk), Delay time: 5 (red star), Delay time: 6 (green circle), Delay time: 7 (cyan diamond), Delay time: 8 (magenta inverted triangle), Delay time: 9 (yellow inverted triangle), Delay time: 10 (blue left-pointing triangle)

FIGURE 5: SIEs when embedding dimension are 7.

Based on the research above, the mutation detection ability of SIE with different parameters combinations (s, m) is analyzed so as to find the rule for the selection of s and m . $x_z(t)$ defined in formula (11) is used as normal signal and $x(t)$ in formula (8) as the failure one. It is obvious that $x(t)$ is the sum of $x_z(t)$ and $x_{fs}(t)$. The addition of $x_{fs}(t)$ inevitably results in the dynamics variation of the signal and also its randomness and complexity. Therefore, SIEs of $x(t)$ and $x_z(t)$ with the same parameters combination are surely different. With certain parameters combination (s, m), the SIEs of the two signals are calculated; then, the D -value of them is calculated and treated as the indicator of SIE's mutation detection ability of corresponding parameter combination. In order to reveal the randomness and complexity variation generated by the addition of $x_{fs}(t)$, we add no noise to the two signals; that is, $n(t) = 0$. In the research, the region of m is 1–7; the value of s is selected based on the following principle: if $m = 1$, the region of s is 2–8; if m is in the scope 2–7, the region of s is $m-8$. Calculate the SIEs of signals $x(t)$ and $x_z(t)$ with determine parameters combination (s, m); name

TABLE 1: Variances of SIEs based on different (s, m) .

Parameters	Values	Embedding dimension						
		1	2	3	4	5	6	7
Partition number	2	0	$1.57e-5$	$7.55e-4$	0.0103	0.006	0	0.032
	3	0	$2.01e-5$	$1.91e-5$	$2.97e-4$	0.0026	0.0037	0.0175
	4	0	$8.40e-6$	$3.28e-5$	$1.98e-5$	$4.44e-5$	0.0018	0.0117
	5	0	$3.56e-6$	$2.19e-5$	$1.72e-6$	$7.04e-6$	$1.15e-5$	0.001
	6	0	$9.61e-7$	$1.34e-5$	$1.61e-6$	$1.47e-5$	$1.52e-5$	0
	7	0	$9.13e-7$	$7.99e-6$	$1.73e-6$	$1.71e-6$	$4.97e-6$	$1.05e-6$
	8	0	$9.87e-7$	$1.39e-7$	$1.62e-6$	$1.27e-6$	$4.55e-6$	$2.59e-7$
	9	0	$1.09e-6$	$1.38e-6$	$1.76e-7$	$8.66e-6$	$4.49e-7$	
	10	0	$9.99e-7$	$1.24e-7$	$2.14e-7$	$1.28e-7$	$6.84e-7$	

TABLE 2: D -values of SIE and SIE_z with different parameters combination.

Parameters	Value	Partition number						
		2	3	4	5	6	7	8
Embedding dimension	1	0.0061	0.0347	-0.0322	-0.0276	-0.0486	0.0354	-0.0299
	2	0.0049	-0.0438	0.0123	0.0042	-0.0458	-0.0213	-0.0356
	3		-0.00014	-0.0227	-0.0087	-0.0451	-0.075	-0.0304
	4			-0.0203	-0.0991	-0.0319	-0.0235	-0.0586
	5				-0.072	-0.0478	-0.0406	-0.1105
	6					-0.1093	-0.0318	-0.0518
	7						-0.069	-0.0892

them as SIE and SIE_z , respectively; then, the D -value SIE – SIE_z is calculated. The D -values with all possible parameters combinations are calculated and shown in Table 2:

$$x_z(t) = x_{is}(t) + n(t). \quad (11)$$

By analyzing the D -values in Table 2, the following conclusions are gained. All the D -values are negative when $m \geq 3$; it illustrates that the addition of $x_{fs}(t)$ results in the increase of regular component and also the decrease of the signal's complexity. However, the D -values have both positive value and negative value when $m < 3$; the reason is analyzed and concluded as if the chosen m value is too small (1 or 2) the method does not work very well; indeed, when m is small, the number of spatial symbol series is small too, so that it is not possible to accurately reflect the dynamic variation. Therefore, $m \geq 3$ is one principle for the selection of embedding dimension. In addition, the increase of s will result in immense increase of operation and computation time, as in the following cases: when parameters combination is (5, 4, 3), the computation time of SIE is 0.253 s; when parameters combination is (8, 5, 3), the time is 42.076 s; and when it is (8, 7, 3), 3010.276 s is used (the operation platform is Matlab 7.11.0 (R2010b); main configuration of the computer is i5-2400 CPU @ 3.1 GHz and 4 G Memory). So, too large value is inadvisable for s . Furthermore, the absolute D -value does not increase with the increase of s and m . So, the preferable region of s and m is 3–8. The parameters combination (5, 4, 3) is selected and used in the following research as the excellent mutation detection ability and the approving operation speed.

3.3. Simulation Signal Analyzing. In this section, a simulation signal used to imitate the fault degradation of pump is designed. The SIE and PE of the simulation signal are extracted; their performance of describing the fault degradation process is checked. The expression of simulation signal is as below:

$$x_r(t) = x_{is}(t) + 0.1t^2 x_{fs}(t) + n(t), \quad (12)$$

where $0.1t^2 x_{fs}(t)$ stands for the fault component and its amplitude $0.1t^2$ is used to represent the degradation degree of pump and the other parameters are the same as $x(t)$. The sampling frequency is 1024 Hz and sampling number is 20480. The data is divided into ten pieces; each piece has 2048 sampling points, and the ten pieces' data are employed to imitate the degradation stages of pump as the fault degree deepens. In order to analyze the influence of noise, different intensity noise is added to the signal; the SNRs of it is 2 dB, 1 dB, -1 dB, -2 dB, -3 dB, and no noise, respectively; Figure 6 is the time domain waveform of it and SNR of it is -1 dB.

Firstly, the signal without noise is analyzed. SIEs of the ten pieces' data are calculated and shown in Figure 7. In order to reveal the degradation indication ability of SIE, the PEs of them are also calculated. Based on the experience of [22], $m = 4$ and $\tau = 3$ are selected as the parameters of PE, and the result is shown in Figure 8. SIEs and PEs of the simulation signal with noise are also calculated and shown in Figures 7 and 8, respectively.

The two figures reflect the degradation indication ability of SIE and PE; from them, we can conclude that both of the

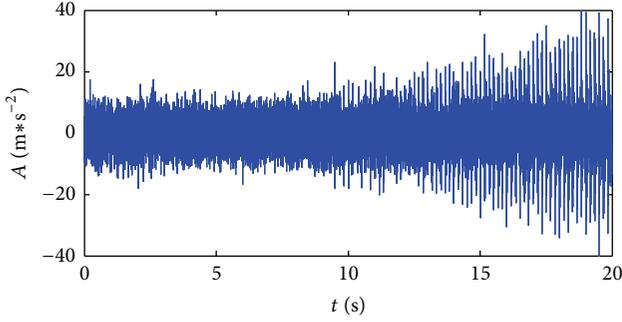


FIGURE 6: Waveform of fault degradation simulation signal.

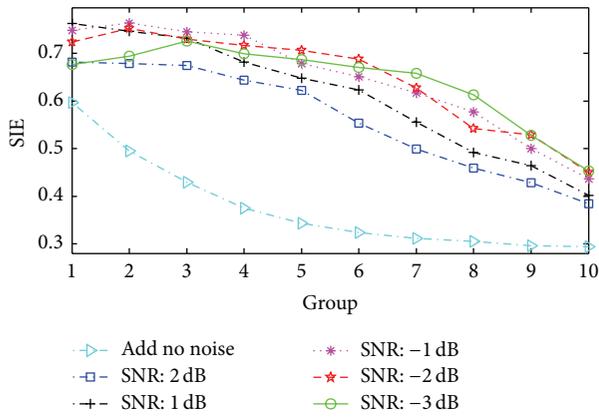


FIGURE 7: SIEs of the fault degradation simulation signal.

indicators present clear descend trend as the fault degradation, which illustrates the reduction of signal's randomness and complexity. The situation fits the principle that the degradation of certain fault results in the increase of regular component and then the decrease of signal's complexity. It is also clear that the decrease scope of SIE is larger than the one of PE no matter what the noise background is. The phenomenon reflects the better fault degradation description ability of SIE. In addition, when the noise intensity is low, the descend trend of both of the indicators is stable. However, when the noise is intense, the two indicators increase at first and then decrease. The reason is analyzed and concluded as at the fault initial stage the fault degradation increases the randomness and complexity of the signal for the intense noise; when the fault develops to a certain degree, the regular component is dominant and the continuous degradation will lead to the decrease of randomness and complexity of the signal.

The following conclusions can be gained based on the above analysis. Both SIE and PE can indicate the fault degradation preferably and SIE has better performance. The noise will weaken the performance of the two indicators when they are describing the degradation trend of faint fault.

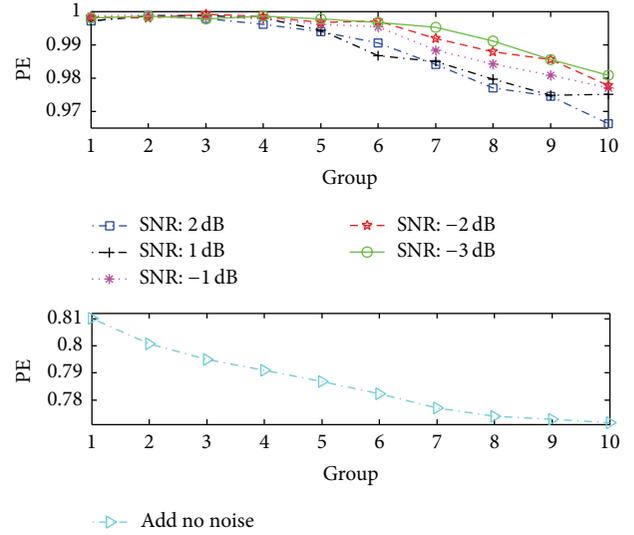


FIGURE 8: PEs of the fault degradation simulation signal.

4. Degradation State Identification Strategy of Hydraulic Pump

There are two steps in the DSI of hydraulic pump; one is the degradation feature extraction and the other is the state identification. SIE and PE are used as degradation features and FCM clustering algorithm is used to identify the degradation state of pump.

4.1. Fuzzy C-Means Algorithm. FCM clustering algorithm is an unsupervised dynamic clustering method, which dims the traditional division definition and adopts membership degree to realize the clustering. Considering a sample set $\{x_1, x_2, \dots, x_n\}$, $x_i \in R$, which is required to be divided into C categories. The aim of FCM clustering algorithm is obtaining each category's clustering center by minimizing the weighted sum of inner-cluster square errors.

The steps of FCM clustering are as follows [8].

(1) Firstly, the vibration signals are analyzed and then the training feature set TS and the testing feature set CS are extracted.

(2) The following parameters are used in the FCM clustering: C ($2 \leq C \leq n$, n is the sample number) is the categories number. ε is the iteration stop threshold. And λ is the smoothing parameter, which controls the sharing degree. $U^{(0)}$ is the initial division matrix. $Z^{(0)}$ is the initial clustering center. The above parameters should be initialized and the iteration counter is $p = 0$.

(3) Based on the training feature set TS, the division matrix $U^{(p)}$ can be updated with

$$\mu_{ij}^{(p)} = \frac{1}{\sum_{k=1}^C \left(g_{ij}^{(p)} / g_{kj}^{(p)} \right)^{2/(\lambda-1)}},$$

$$1 \leq i \leq C, 1 \leq j \leq n, \forall i, j, g_{ij}^{(p)} > 0,$$

$$\begin{aligned} \mu_{ij}^{(p)} &= 1, \quad \exists i, j, g_{ij}^{(p)} = 0, \\ \mu_{ij}^{(p)} &= 0, \quad \exists i, j, g_{ij}^{(p)} = 0, k \neq j, \end{aligned} \quad (13)$$

where μ_{ij} is the membership degree of sample j ($j = 1, 2, \dots, n$) that belongs to category i ($i = 1, 2, \dots, C$) and its requirements are $\mu_{ij} \in [0, 1], \forall i, j; \sum_{i=1}^C \mu_{ij} = 1, \forall j; 0 < \sum_{j=1}^n \mu_{ij} < n, \forall i$. g_{ij} represents the distance between x_j and clustering center z_i , which is often measured in Euclidean space:

$$g_{ij}(x_j, z_i) = \|x_j - z_i\|. \quad (14)$$

(4) The clustering center can be updated with

$$z_i^{(p+1)} = \frac{\sum_{j=1}^n (\mu_{ij}^{(p+1)})^\lambda x_j}{\sum_{j=1}^n (\mu_{ij}^{(p+1)})^\lambda}, \quad 1 \leq i \leq C. \quad (15)$$

(5) If $\|Z^{(p)} - Z^{(p+1)}\| < \varepsilon$, the iteration stops. Otherwise, let $p = p + 1$ and turn to step (2). Through iterative optimization, the division matrix U and clustering center matrix Z can be obtained.

(6) Then, the testing feature set CS is used to realize the state identification.

4.2. Strategy. The DSI is shown in Figure 9. Training samples are de-noised at first. Then, SIE and PE of them are extracted and the degradation feature vectors are united with them. The FCM model is trained with training feature vectors and the clustering centers of each degradation state are gained. Test samples are preprocessed and the test feature vectors are gained. Finally, the degradation state is identified and the availability of the proposed method is demonstrated.

5. Experiment Validation

5.1. Experiment Rig. The practical vibration signals of pump are collected from the test bench [25] which is shown in Figure 10. The type of pump is SY-10MCY14-1EL. The driving motor's type is Y132M-4 and its rated speed is 1480 r/min. The piezoelectric acceleration sensor of type CA-YD-139 is selected and rigidly connected to the pump's end cover. We adopt dynamic signal test and analysis system of type DH-5920 to collect the vibration signal. Loose boot and sliding boot wear are the typical fault patterns of pump; therefore, the vibration signal of pump failure with single loose boot or single sliding boot wear is analyzed in this paper. In order to obtain more real vibration signal, the failure plungers abandoned in the examination of equipment are used to replace the normal one. Five loose boot plungers with different degree and four sliding boot wear plungers with different degree are selected; parts of the plungers are shown in Figure 11. In order to determine the fault degree of loose boot, the loose degree is defined as the largest radial shift distance between the boot and the plunger. Vernier caliper is used to measure the loose degree, and the result is 0.12 mm,

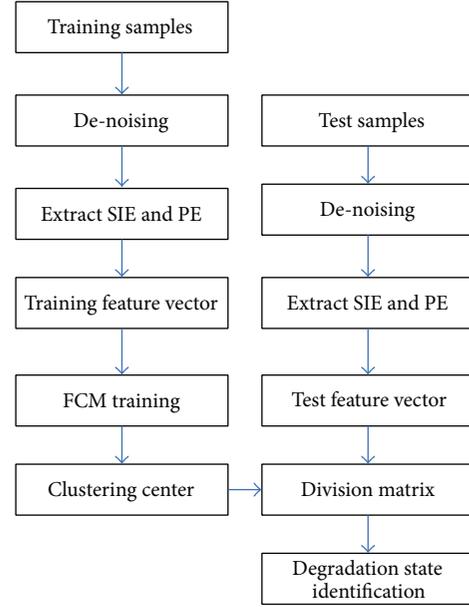


FIGURE 9: DSI strategy.

0.18 mm, 0.3 mm, 0.42 mm, and 0.64 mm. In the same way, the wearing degree is defined as the increment of boot's edge diameter. Because the sliding boot wearing will result in the increase of boot's edge diameter, the wearing degree is used to determine the fault degree of sliding boot wear. The wearing degrees of the four plungers are 0.1 mm, 0.16 mm, 0.26 mm, and 0.48 mm, respectively. In addition, the normal state is considered as the special case of fault state. Two hundred groups' data are collected for each state, and each group has 2048 sampling points; the sampling frequency is 50 KHz and sampling interval is 30 s. The pressure of the main relief valve is 10 MPa and the motor speed is its rated speed.

Parts of the collected vibration signals are shown in Figure 12.

5.2. Extraction of Degradation Feature. As the noise influences the performance of SIE and PE, it is necessary to filter noise before extracting SIE and PE; in this paper, the CNC de-noising method proposed in [25] is used to preprocess the collected vibration signals. In order to check the real fault degradation indication ability of SIE and PE, ten samples of each state are randomly selected; then, SIE and PE of them are calculated and shown in Figures 13 and 14.

It can be seen that SIE and PE of normal samples are the largest, which illustrates the randomness and complexity of normal signal are the largest. Along with the degradation of fault, no matter which fault pattern, both SIE and PE decrease obviously; the phenomenon illustrates that the degradation of certain fault leads to the increase of regular component then the decrease of randomness and complexity of signal. For the same fault, the SIE D -value between different degradation states is larger than the one of PE; the result reflects the better performance of SIE. For sliding boot wear samples, both SIE and PE distinguish the degradation states clearly. With regard to loose boot samples, SIE also shows satisfactory

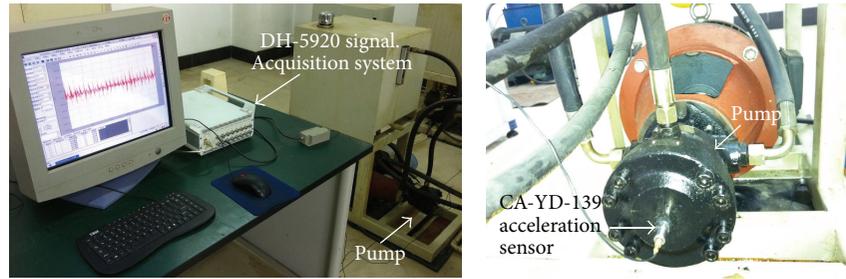


FIGURE 10: Test bench of hydraulic pump.



FIGURE 11: Part plungers used in the experiment.

performance; however, several PEs of normal samples are smaller than the one of corresponding fault samples; the case illustrates that SIE is more excellent than PE when the fault is faint.

5.3. Identification of Degradation State. In this section, the degradation state of pump is identified with FCM clustering algorithm. The selection of smoothing parameter λ is important as it regulates the sharing degree of different degradation states; inappropriate λ will affect the result. As to obtain suitable λ , numerous trials are carried out on the degradation samples of pump; the relationship between λ and correct clustering ratio is shown in Figure 15. The correct clustering ratio reaches its maximum when $\lambda = 3$. The following are the training of FCM clustering model and the identification of degradation state when $\lambda = 3$.

With consideration of sliding boot wear fault, the training set contains 250 samples which are randomly selected from the sample sets of 5 fault degrees, 50 groups in each set. And another 250 groups are selected as test set in the same way. The SIE and PE of training samples are calculated and a 250×2 training feature set TS_s is obtained. Then, the FCM model is trained with TS_s ; the clustering number is 5. Clustering center matrix Z_s and division matrix U_s are obtained, and Z_s is shown in Table 3.

Calculate the SIE and PE of test set, and the test feature set CS_s is obtained. The division matrix of test set is calculated based on Z_s and CS_s . The degradation state is identified based on the membership degree. The division matrix of normal samples is shown in Table 4; all of the 50 samples

TABLE 3: Clustering centers of sliding boot wear feature set.

Degradation features	Wearing degree (mm)				
	Normal	0.1	0.16	0.26	0.48
PE	0.9975	0.9932	0.9873	0.9826	0.9469
SIE	0.2908	0.2331	0.1948	0.1556	0.1249

are correctly diagnosed as normal state based on the principle of maximum membership degree. The samples of other four degradation states are also used to check the FCM model and the right ratio is also 100%. The identification result is shown in Figure 16.

With consideration of the loose boot fault, the training feature set TS_L and test feature set CS_L are extracted in the same way. The FCM clustering model is trained with TS_L , and the clustering centers are obtained and shown in Table 5.

The degradation states are identified based on Z_L and CS_L , and the result is shown in Table 6 and Figure 17. The total identification ratio is 97.67%.

The practical analysis demonstrates not only the rationality of SIE but also its effectiveness and superiority of using as the degradation feature of pump. In addition, the rationality and effectiveness of adopting FCM clustering algorithm to identify the degradation states of pump are proved.

6. Conclusions

This paper addresses pump degradation identification with the aim of avoiding unexpected failure of it which will result

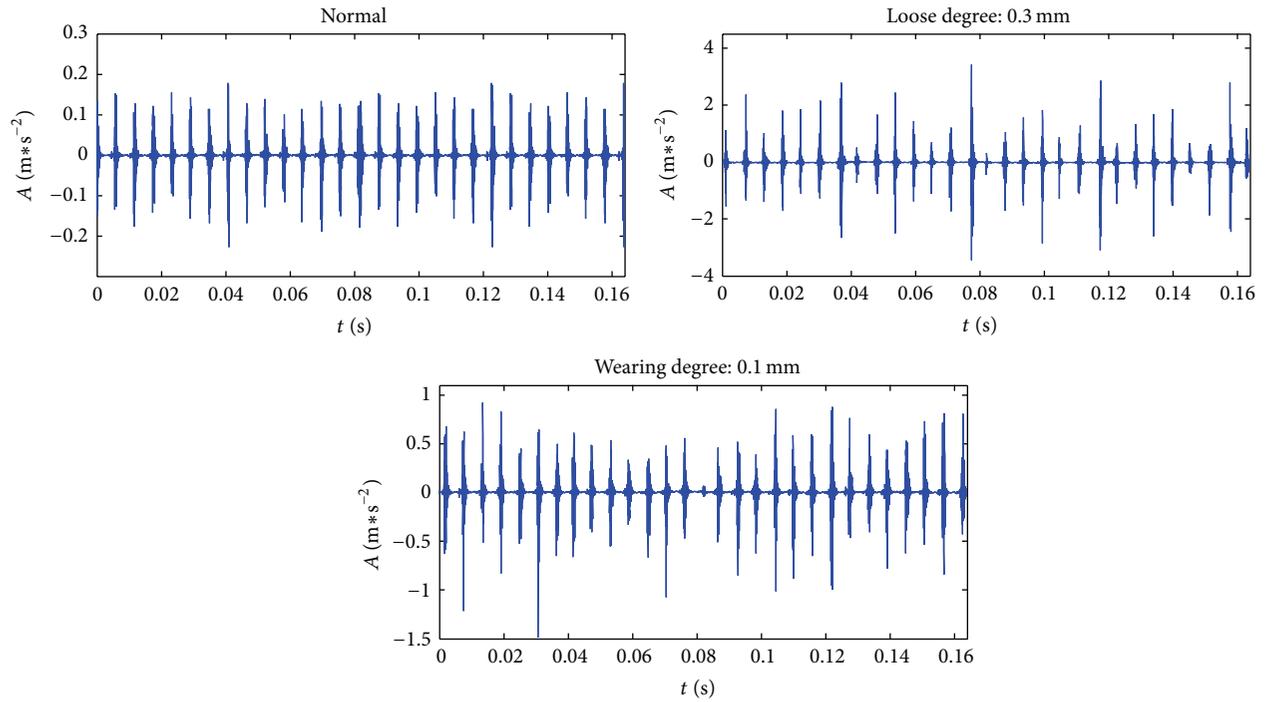


FIGURE 12: Waveform of collected vibration signals.

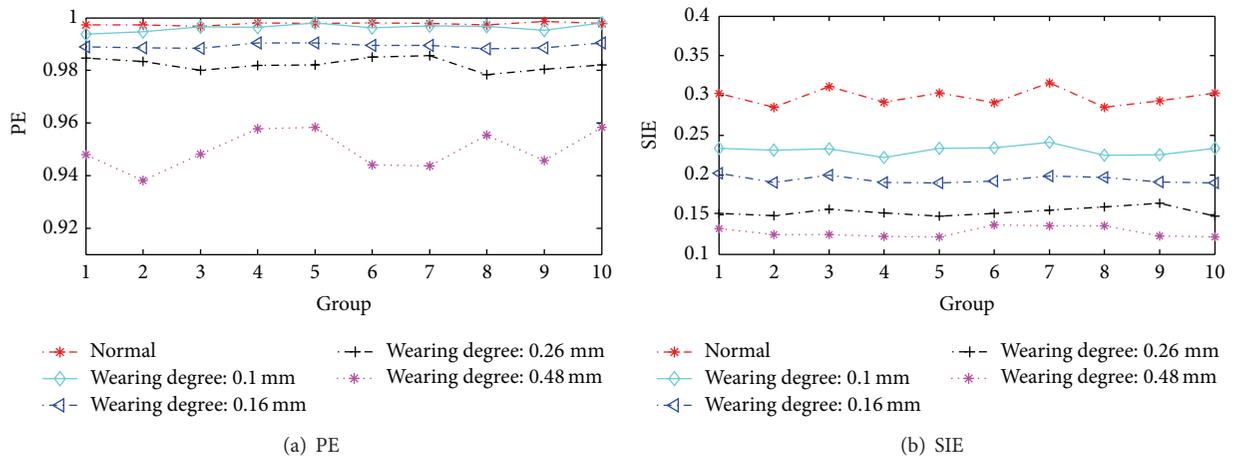


FIGURE 13: SIE and PE of sliding boot wear samples.

TABLE 4: State identification result of normal samples.

Samples	1	2	3	4	...	48	49	50	
Results	Normal	Normal	Normal	Normal	...	Normal	Normal	Normal	
Wearing degree	Normal	0.9973	0.8397	0.8293	0.9454	...	0.98	0.881	0.8696
	0.1 mm	0.0016	0.0835	0.0884	0.0354	...	0.0126	0.0636	0.0692
	0.16 mm	$6.2e-4$	0.0393	0.042	0.0109	...	0.0041	0.0287	0.0316
	0.26 mm	$3.1e-4$	0.0225	0.0242	0.0052	...	0.002	0.0161	0.0178
	0.48 mm	$1.7e-4$	0.015	0.0161	0.0031	...	0.0013	0.0106	0.0118

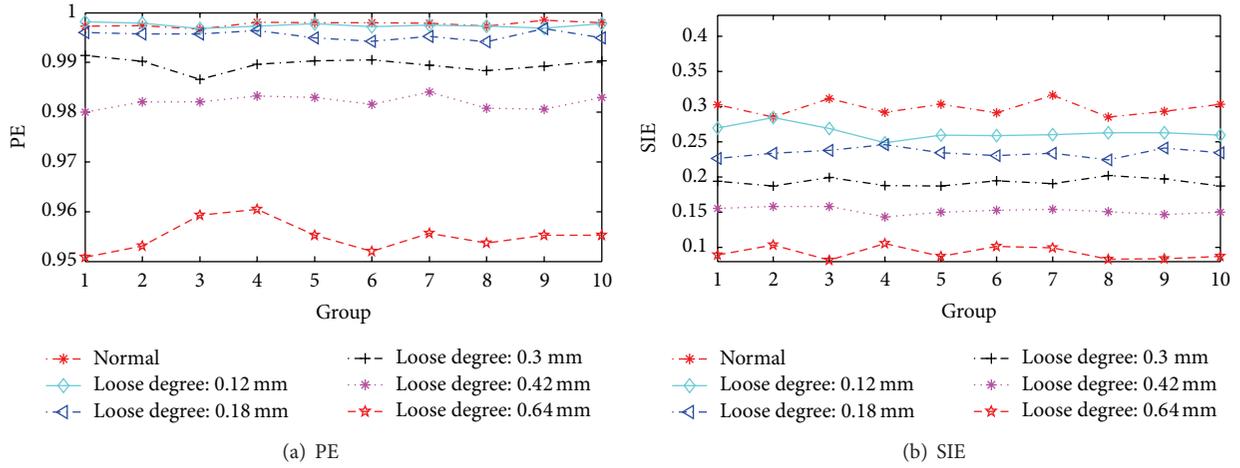


FIGURE 14: SIE and PE of loose boot samples.

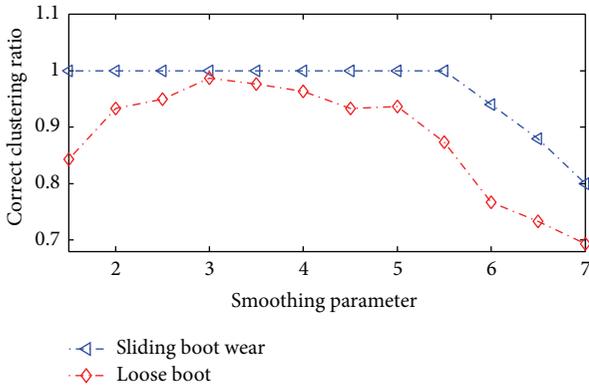


FIGURE 15: Relationship between λ and correct clustering ratio.

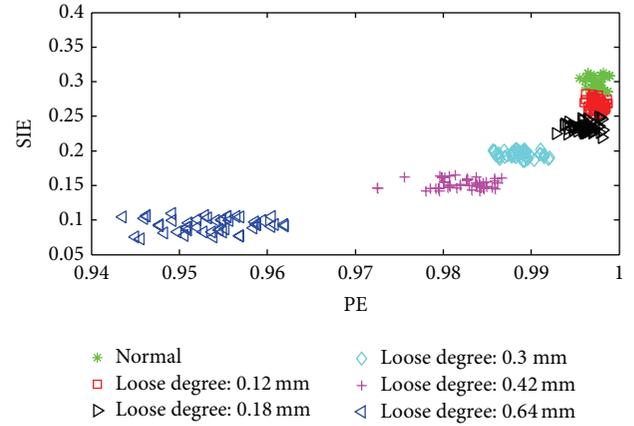


FIGURE 17: Identification result of loose boot samples.

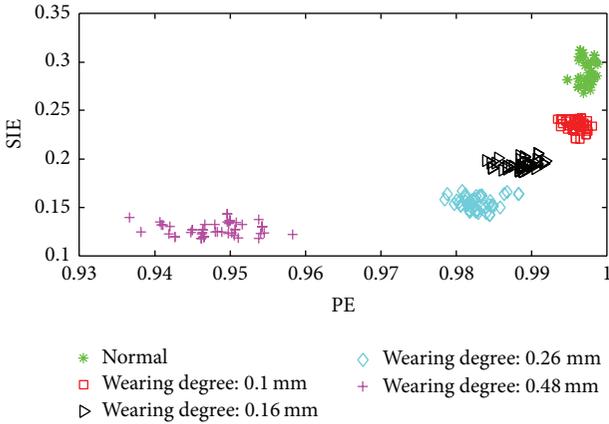


FIGURE 16: Identification result of sliding boot wear samples.

in loss of production time. PE is studied and regarded as the degradation feature of pump. In order to reflect the dynamics variation of pump's vibration signal more excellently in detail, SIE is proposed. Different parameters which affect its performance have been discussed, and it has been concluded

TABLE 5: Clustering centers of loose boot feature set.

Degradation features	Loose degree (mm)					
	Normal	0.12	0.18	0.3	0.42	0.64
PE	0.9974	0.9966	0.9953	0.9886	0.9842	0.9549
SIE	0.289	0.2656	0.2346	0.1938	0.15	0.0951

TABLE 6: Identification result of loose boot test samples.

Loose degree (mm)	Normal	0.12	0.18	0.3	0.42	0.64
Identification ratio (%)	92	100	94	100	100	100

that embedding dimension should be in the range of 3–8 and partition number should be equal or greater than embedding dimension, based on which delay time can be randomly selected in the range of 1–10. The simulation analysis results reveal the excellent performance of SIE. Both SIE and PE are used as the degradation features of pump, and FCM clustering algorithm is adopted to identify the degradation

states. Experimental research on vibration signals of pump has found that both SIE and PE could effectively indicate the dynamic change of them. The degradation state identification result proved the rationality and effectiveness of proposed method.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (no. 51275524) and the General Armaments Equipment Department Support Research Project.

References

- [1] Z. Zhao, M. Jia, F. Wang, and S. Wang, "Intermittent chaos and sliding window symbol sequence statistics-based early fault diagnosis for hydraulic pump on hydraulic tube tester," *Mechanical Systems and Signal Processing*, vol. 23, no. 5, pp. 1573–1585, 2009.
- [2] J. Du, S. P. Wang, and H. Y. Zhang, "Layered clustering multi-fault diagnosis for hydraulic piston pump," *Mechanical Systems and Signal Processing*, vol. 36, no. 2, pp. 487–504, 2013.
- [3] J. P. Wang and H. T. Hu, "Vibration-based fault diagnosis of pump using fuzzy technique," *Measurement*, vol. 39, no. 2, pp. 176–185, 2006.
- [4] H. X. Chen, P. S. K. Chua, and G. H. Lim, "Fault degradation assessment of water hydraulic motor by impulse vibration signal with Wavelet Packet Analysis and Kolmogorov-Smirnov Test," *Mechanical Systems and Signal Processing*, vol. 22, no. 7, pp. 1670–1684, 2008.
- [5] A. K. S. Jardine, D. Lin, and D. Banjevic, "A review on machinery diagnostics and prognostics implementing condition-based maintenance," *Mechanical Systems and Signal Processing*, vol. 20, no. 7, pp. 1483–1510, 2006.
- [6] A. J. Yan, W. X. Wang, C. X. Zhang, and H. Zhao, "A fault prediction method that uses improved case-based reasoning to continuously predict the status of a shaft furnace," *Information Sciences*, vol. 259, no. 20, pp. 269–281, 2014.
- [7] J. Liu, W. Wang, and F. Golnaraghi, "A multi-step predictor with a variable input pattern for system state forecasting," *Mechanical Systems and Signal Processing*, vol. 23, no. 5, pp. 1586–1599, 2009.
- [8] Y. Pan, J. Chen, and X. L. Li, "Bearing performance degradation assessment based on lifting wavelet packet decomposition and fuzzy c-means," *Mechanical Systems and Signal Processing*, vol. 24, no. 2, pp. 559–566, 2010.
- [9] Y. N. Pan, J. Chen, and L. Guo, "Robust bearing performance degradation assessment method based on improved wavelet packet-support vector data description," *Mechanical Systems and Signal Processing*, vol. 23, no. 3, pp. 669–681, 2009.
- [10] J.-B. Yu, "Bearing performance degradation assessment using locality preserving projections," *Expert Systems with Applications*, vol. 38, no. 6, pp. 7440–7450, 2011.
- [11] C. Hua, G.-H. Xu, Q. Zhang, Y.-Z. Zhang, J. Xie, and S.-Z. Li, "Long-term potential performance degradation analysis method based on dynamical probability model," *Expert Systems with Applications*, vol. 39, no. 4, pp. 4410–4417, 2012.
- [12] A. Widodo and B. S. Yang, "Machine health prognostics using survival probability and support vector machine," *Expert Systems with Applications*, vol. 38, no. 7, pp. 8430–8437, 2011.
- [13] S. J. Dong, B. P. Tang, and R. X. Chen, "Bearing running state recognition based on non-extensive wavelet feature scale entropy and support vector machine," *Measurement*, vol. 46, no. 10, pp. 4189–4199, 2013.
- [14] H. T. Pham and B.-S. Yang, "Estimation and forecasting of machine health condition using ARMA/GARCH model," *Mechanical Systems and Signal Processing*, vol. 24, no. 2, pp. 546–558, 2010.
- [15] T. Benkedjouh, K. Medjaher, N. Zerhouni, and S. Rechak, "Remaining useful life estimation based on nonlinear feature reduction and support vector regression," *Engineering Applications of Artificial Intelligence*, vol. 26, no. 7, pp. 1751–1760, 2013.
- [16] M. Costa, A. L. Goldberger, and C.-K. Peng, "Multi-scale entropy analysis of biological signals," *Physical Review E*, vol. 71, no. 2, Article ID 021906, 18 pages, 2005.
- [17] S.-D. Wu, C.-W. Wu, K.-Y. Lee, and S.-G. Lin, "Modified multiscale entropy for short-term time series analysis," *Physica A: Statistical Mechanics and Its Applications*, vol. 392, no. 24, pp. 5865–5873, 2013.
- [18] C. Bandt and B. Pompe, "Permutation entropy: a natural complexity measure for time series," *Physical Review Letters*, vol. 88, no. 17, Article ID 174102, 2002.
- [19] C. Bandt, "Ordinal time series analysis," *Ecological Modelling*, vol. 182, no. 3-4, pp. 229–238, 2005.
- [20] N. Nicolaou and J. Georgiou, "Detection of epileptic electroencephalogram based on Permutation Entropy and Support Vector Machines," *Expert Systems with Applications*, vol. 39, no. 1, pp. 202–209, 2012.
- [21] B. Frank, B. Pompe, U. Schneider, and D. Hoyer, "Permutation entropy improves fetal behavioural state classification based on heart rate analysis from biomagnetic recordings in near term fetuses," *Medical and Biological Engineering and Computing*, vol. 44, no. 3, pp. 179–187, 2006.
- [22] R. Q. Yan, Y. B. Liu, and R. X. Gao, "Permutation entropy: a nonlinear statistical measure for status characterization of rotary machines," *Mechanical Systems and Signal Processing*, vol. 29, no. 5, pp. 474–484, 2012.
- [23] G. Ouyang, J. Li, X. Liu, and X. Li, "Dynamic characteristics of absence EEG recordings with multiscale permutation entropy analysis," *Epilepsy Research*, vol. 104, no. 3, pp. 246–252, 2013.
- [24] W. Pedrycz, A. Amato, V. di Lecce, and V. Piuri, "Fuzzy clustering with partial supervision in organization and classification of digital images," *IEEE Transactions on Fuzzy Systems*, vol. 16, no. 4, pp. 1008–1026, 2008.
- [25] Y. K. Wang, H. R. Li, and P. Ye, "Fault feature extraction of hydraulic pump based on CNC De-noising and HHT," *Journal of Failure Analysis and Prevention*, vol. 15, no. 1, pp. 139–151, 2015.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

